On the Contribution of Human Capital to Growth: Theory, Identification, and Evidence

Kahn, James and Jong-Soo Lim

Working Paper No. 377
April 1994

On the Contribution of Human Capital to Growth: Theory, Identification, and Evidence

James A. Kahn and Jong-Soo Lim

Rochester Center for Economic Research
Working Paper No. 377

On the Contribution of Human Capital to Growth: Theory, Identification, and Evidence

James A. Kahn
and
Jong-Soo Lim*
Department of Economics
University of Rochester

April 1994

*The authors thank Mark Bils, Bruce Hansen, Changyong Rhee, and workshop participants at the University of Rochester and the University of Pennsylvania. Any remaining errors are the authors'.

Abstract

This paper quantifies the relative importance of skilled–labor augmenting technical change and general technical change. We develop a general equilibrium model in a multi-industry setting where skilled labor–augmenting technological progress can be distinguished from other sources of economic growth because of variation in skilled labor's share across industries. The results, based on a panel of 21 U.S. manufacturing industries, indicate that skilled labor–augmenting technological progress is the significant factor in productivity growth. Growth in conventional total factor productivity vanishes once the role of skilled labor and the growth in its human capital are properly accounted for.

1 Introduction

This paper examines the importance of skilled labor in the cross–sectional and aggregate variation of output growth. Many recent studies have documented the role of human capital in accounting for growth in cross-country data.[1] In those studies, the lack of consistent measurement of human capital poses a major hurdle to researchers. In constrast, we use cross–industry data within the United States, where the skill level of workers of given educational experience is plausibly similar, and where other factors that may influence growth, such as political stability, social institutions, and the like, are held constant. Thus our analysis is more closely related to the recent literature on the growth of the wage differential between skilled and unskilled labor, and the measurement of "skilled–labor–augmenting technical change."

Disentangling different sources of growth is a fundamental problem in time series data. Consider a production function: $Y = A \cdot F(BK, HN)$ where A, B, and H are technology variables. We can think of A as any disembodied technological progress that is external to individual decisions. B represents quality improvements in capital. With conventional growth accounting, appropriate measurement of capital should take account of changes in B, so that in fact capital would be measured not as K, but as BK. We can think of H as the average skill level of a worker. While A and H are conceptually distinct, there is a fundamental difficulty in disentangling their contributions to output from time series data, due to the fact that we only measure inputs to the formation of H, such as years of schooling. First, it is likely that a given level of conventionally measured inputs give rise to changing levels of H, over time. *Ceteris paribus,* sixteen years of schooling surely results in a higher level of H today than it did 50 years ago. Knowledge accumulates, and normally one would expect that each generation can attain a given level of knowledge with fewer years of schooling. Thus, for example, if the production function for H were $H_t = H_0 e^{gt} Z_t^{\gamma}$, where Z_t is the input, we would need to know both g and γ in order to isolate the contribution of growth in H to output growth. The parameter g cannot be identified separately from trend growth in A with time series data. Moreover, although in theory γ is identified, time series variation in Z is likely to be so small and gradual as to make precise estimation of γ impossible.

---

[1] See, for example, Barro (1991) and Barro and Sala-i-Martin (1989).

In this paper we surmount these difficulties by exploiting cross–sectional variation in measured inputs. The basic idea is that the accumulation of knowledge affects labor differently according to what we will call "skill level." Improvements in electronics and computers, for example, presumably have a larger impact on the effective labor input of engineers and statisticians than of farm workers and janitors. If we can identify the extent to which industries differ in skilled relative to unskilled labor input, we can use that cross–sectional variation to identify the contribution of this phenomenon to overall growth.

Before embarking on that task, a natural first check would be to see whether productivity growth is higher in industries with higher shares of skilled labor. Although our analysis is consistent with any correlation, it is unlikely that we would identify skilled labor and knowledge as important factor in growth if this correlation were zero or negative. Figure 1 shows the scatter diagram of conventional Solow residuals and the share of skilled labor of 21 U.S. manufacturing industries. Skilled labor's share is measured in two different ways:[2] In the first panel (Figure 1a) it is the share of workers with 12 or more years of education; in the second panel (Figure 1b) it is the share of non–production workers. In both cases a clear pattern of positive correlation emerges, though with considerable variation, and with a number of outliers in both directions.

Given the assumption that we can categorize labor as "skilled" and "unskilled," we can test the hypothesis that the productivity of skilled and unskilled labor are affected differently by technological progress and increases in knowledge, and thereby assess the importance of this factor in the growth of total factor productivity (TFP henceforth). We do this by quantifying the extent to which productivity growth in a cross–section of U.S. manufacturing industries is tied to these industries' shares of skilled labor inputs. We identify the portion of total factor productivity growth that we can associate with skilled labor's share as due to the accumulation of skills, or as "skilled labor–augmenting technological progress," while the remainder we term "generalized" productivity growth. We find strong evidence of high growth of such skills in the U.S. manufacturing sector, leading to the conclusion that the growth in human capital plays a significant role in explaining output growth. This evidence confirms the findings of Barro (1991), Barro

[2] We discuss these measures in more detail below and in the Appendix.

and Lee (1993), and Levine and Renelt (1991). Furthermore, we find that this skill factor accounts for essentially 100 percent of conventional total factor productivity growth.[3]

In addition to exploring the empirical relationship between skilled labor, output growth, and productivity growth, we develop a simple equilibrium growth model to explore other steady state implications for disaggregated data. The main implication of the model is that the differences in the long run growth rates of output and relative prices can be characterized as functions of the skilled labor shares of respective industries. The data also provide qualitative support for this aspect of the model.

Our findings have important implications for growth models. First, they refine the neoclassical approach by endogenizing total factor productivity growth in a quantifiable way. Second, they restrict the potential scope of Marshallian externalities in accounting for growth. Lucas (1988) and Romer (1986), among others, have suggested that, theoretically, an aggregate externality can play an important role in providing source of economic growth. In an empirical study, Caballero and Lyons (1989) argue that there is a strong evidence of such externalities at a high level of aggregation of U.S. economy. Our results do not support this view. We argue that once the skill level of labor is properly accounted for, there is no evidence for aggregate externalities through the production function except those that could enter through the each industry's employment of skilled labor.

The organization of this paper is as follows. Section 2 describes the model; Section 3 derives the steady state implication of the model, and Section 4 presents the estimation strategy and empirical evidence. Section 5 discuss possible existence of productive externality in human capital, and section 6 concludes.

2 The Model

2.1 Technology

We distinguish "knowledge" from "human capital" as follows: Knowledge refers to the potential quantity of human capital available to a worker. Only

[3] Gort and Wall (1993) report a similar result.

workers who are "skilled" actually incorporate that knowledge into their labor and attain that level of human capital. This does not mean that the workers we call "unskilled" literally have no skills, only that their human capital does not grow with that of skilled workers. In other words, an advance in knowledge will have a direct impact on the effective labor input of a skilled worker, but only an indirect impact (through growth in other factors of production) on the productivity of unskilled labor. Such an advance corresponds to what has been referred to elsewhere as "skilled labor–augmenting technical change." General technological progress would be equivalent to adding to the effective labor of both skilled and unskilled alike.

Although it is easy to write down a model that endogenizes the growth of knowledge over time, it suffices for the purposes of this paper to let it grow exogenously. We denote the level of knowledge in the economy at date t by the variable H_t. Also for the sake of simplicity we assume that all skilled workers have identical human capital of $H_t > 1$, while unskilled labor has a constant level of human capital normalized to one.[4]

There are M consumption goods produced in industries $1, ..., M$, with physical capital produced in industry 0. Each of the $M + 1$ sectors uses physical capital, unskilled labor, and skilled labor, to produce its output. A representative firm in industry i then has the following technology:

$$Y_{it} = A_{it} F_i(K_{it}, N^s_{it} H_t, N^u_{it}), \tag{1}$$

where K_{it} is physical capital, and N^u_{it} is the number of unskilled workers at the firm.[5] The firm faces market wages W^s_t and W^u_t, an interest rate R_t, and market prices $\{P_{it},\ i = 1, ..., M\}$.

2.2 Preferences

We model the consumer side of the economy only in terms of utility for final goods; we take the characteristics of "skilled" and "unskilled," as well as the growth of knowledge and human capital, as exogenous. This is obviously not

[4] This can be generalized to allowing an exogenous constant growth rate to the efficiency of unskilled labor. This effect was not empirically detectable, though it is econometrically identified.

[5] We classified workers into these two category for expositional simplicity. However, classifying workers into multiple categories can be done just as easily.

necessary, but nothing is gained for the purposes of this paper from modeling the process by which these variables get determined. The empirical analysis can proceed conditional on the observed values.

The economy at time t consists of N_t infinitely–lived agents with identical preferences over M distinct consumption goods in each period of their lives. Each agent maximizes lifetime utility

$$\sum_{t=0}^{\infty} \beta^t U(\tilde{C}_t) : R_+^{M \times \infty} \rightarrow R \tag{2}$$

where $0 < \beta < 1$ is a subjective time discount factor, and $\tilde{C}_t$ is a $(M \times 1)$ consumption vector. Each agent is endowed with one unit of time in each period that can only be used for work effort. A subset N_t^s of the workers are skilled, and consequently have an effective labor supply of H_t; the remaining $N_t^u = N_t - N_t^s$ are unskilled. We assume the efficiency level of the labor of unskilled workers is constant, and set it equal to one.[6] Since workers are paid for their total amount of labor in efficiency units, skilled workers earn labor income of $W_t^s H_t$, where W_t^s is wage of skilled worker per efficiency labor unit at time t, while unskilled workers' labor income at time t would be W_t^u, where W_t^u is defined to be the wage rate for one unit of unskilled labor.

Agents are able to accumulate interest–bearing assets (i.e. physical capital, taken to be good 0, and the numéraire) to help smooth consumption across time. Given a set of prices $\{P_{it},\ i = 1, ..., M\}$, a representative skilled agent's asset holding at the beginning of period $t + 1$ will be

$$\sum_{i=0}^{M} X_{it+1} = \sum_{i=0}^{M} R_{it} X_{it} + W_t^s H_t - \sum_{i=1}^{M} P_{it} C_{it} \tag{3}$$

where R_{it} denotes the real return to the agent's capital holdings X_{it} in industry i at time t. A representative unskilled worker's asset holdings are

$$\sum_{i=0}^{M} X_{it+1} = \sum_{i=0}^{M} R_{it} X_{it} + W_t^u - \sum_{i=1}^{M} P_{it} C_{it}. \tag{4}$$

[6]This can be generalized to allowing an exogenous constant growth rate to the efficiency of unskilled labor. This effect was not empirically detectable, though it is econometrically identified.

There is no uncertainty in the model, and physical capital is homogeneous, hence it has a single price . To summarize the agents decision–making process: Agents, both skilled and unskilled, decide which industry to work in. At the end of the period, after working and obtaining capital and labor income, they decide how much to save and to consume. Therefore, the maximization problem faced by an agent at the beginning of a period is represented by the following dynamic programming problem:

$$V^s(\tilde{X}_t, H_t) = \underset{C,X,}{Max}\ U(\tilde{C}_t) + \beta V^u(\tilde{X}_{t+1}, H_{t+1}) \quad \text{s. t. (3)} \tag{5}$$

if skilled, or

$$V^u(\tilde{X}_t, H_t) = \underset{C,X}{Max}\ U(\tilde{C}_t) + \beta V^u(\tilde{X}_{t+1}, H_{t+1}) \quad \text{s. t. (4)}$$

if unskilled, where X_t is the total value of asset holding.

2.3 Equilibrium

The $M+1$ firms face the following myopic optimization problem:

$$\underset{K_{ji}, N^s_{ji}, N^u_{it}}{\text{Max}} \quad P_{it} A_{it} F(K_{it}, H_t N^s_{it} \ell_{it}, N^u_{it}) - R_{it} K_{it} - W^s_{it} H_{it} N^s_{it} \ell_{it} - W^u_{it} N^u_{it} \tag{6}$$

Firms' optimality condition and zero profit conditions yield that the payment to each input factors must be equal to their marginal revenue products. Thus we have,

$$\left.\begin{array}{l} P_{it} F_1(K_{it}, H_t N^s_{it} \ell_{it}, N^u_{it}) = R_{it} \\ P_{it} F_2(K_{it}, H_t N^s_{it} \ell_{it}, N^u_{it}) = W^s_t \\ P_{it} F_3(K_{it}, H_t N^s_{it} \ell_{it}, N^u_{it}) = W^u_t \end{array}\right\} \forall i = 0, 1, ... M \tag{7}$$

Since there is no uncertainty in the model, it will of course be the case that $R_{it} = R_{jt}\ \forall i, j$.

The solution to the households' dynamic programming problem can be characterized by (3), (4), (5), which give rise to the usual optimality conditions:

$$\frac{U_i(\tilde{C}_t)}{P_{it}} = \frac{U_j(\tilde{C}_t)}{P_{jt}} \quad i, j = 1, 2, ..., M \tag{8}$$

$$\frac{U_i(\tilde{C}_t)}{P_{it}} = \beta R_{it+1} \frac{U_i(\tilde{C}_{t+1})}{P_{it+1}} \ i = 1, ..., M. \tag{9}$$

3 Steady State

With specific assumptions on $U(\cdot)$ and $F(\cdot)$, we can describe the properties of a balanced growth steady state. This is useful for two reasons: First, the model may be more applicable to longer–term averages than to higher frequency data (which may be more subject to measurement error, for example); second, the equilibrium analysis involved in analyzing a steady state draws out implications for other variables, notably relative prices.

For the steady state analysis only, we assume constant elasticity preferences and technology. The utility function of a household is assumed to have the following form:

$$U(\tilde{C}_t) = \left[\sum_{i=1}^{M} \sigma_i C_{it}^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \tag{10}$$

A firm combines physical capital and labor to produce a tangible output. Production functions of industry $0, 1, 2, ..., M$ take the following form:

$$y_{it} = A_{it}[\alpha_K K_{it}^{1-\theta} + \alpha_u N_{it}^{u^{1-\theta}} + \alpha_s (H_t N_{it}^s)^{1-\theta}]^{\frac{1}{1-\theta}},\ i = 0, 1, .., M. \tag{11}$$

Cobb–Douglas production is the special case in which $\theta = 1$. The population N grows geometrically at rate n.

We are looking for a steady state in which R is constant, and in which there is a constant savings rate as a percentage of total income. We will let σ_0 denote that rate. We can assume without loss of generality that $\sum_{i=0}^{M} \sigma_i = 1$. We also suppose that the output of each sector, prices, capital, and labor inputs all grow at constant rates. Given the form of the utility function, we know that $\forall i, j = 1, ...M,\ P_{it} y_{it}^{\gamma} / \sigma_i = P_{jt} y_{jt}^{\gamma} / \sigma_j$. Now let variables with a "^" refer to steady state growth rates. Then we know that $\hat{P}_i + \gamma \hat{y}_i = \hat{P}_j + \gamma \hat{y}_j\ \forall i, j$. We also know that given (7) and (11), we have

$$\begin{aligned} \alpha_K P_{it} (y_{it}/K_{it})^{\theta} A_{it}^{1-\theta} &= R, \\ \alpha_s P_{it} (y_{it}/N_{it}^s)^{\theta} A_{it}^{1-\theta} &= W^u \\ \alpha_s P_{it} (y_{it}/N_{it}^s)^{\theta} (A_{it} H)^{1-\theta} &= W^s H \end{aligned} \tag{12}$$

or, in terms of steady state growth rates,

$$\begin{aligned} \hat{P}_i + \theta(\hat{y}_i - \hat{K}_i) + (1-\theta)\hat{A}_i &= 0, \\ \hat{P}_i + \theta(\hat{y}_i - \hat{N}_i^u) + (1-\theta)\hat{A}_i &= \hat{W}^u, \\ \hat{P}_i + \theta(\hat{y}_i - \hat{N}_i^s) + (1-\theta)(\hat{H} + \hat{A}_i) &= \hat{W}^s + \hat{H}. \end{aligned} \tag{13}$$

There is no exact steady–state log–linear equation for the production function unless $\theta = 1$ (the Cobb–Douglas case) because otherwise the factor shares are not constant. We can approximate this relationship if θ is not very far from one, though, by using the average shares for each industry, which are denoted by $\bar{\alpha}_j,\ j = K, u, s$:

$$\hat{y}_i = \hat{A}_i + \bar{\alpha}_{Ki}\hat{K}_i + \bar{\alpha}_{ui}\hat{N}_i^u + \bar{\alpha}_{si}(\hat{N}_i^s + \hat{H}), \qquad i = 0, 1, ..., M, \tag{14}$$

Finally, from the consumer's first–order conditions, we have

$$\hat{P}_i + \gamma\hat{y}_i = \text{ constant}, \qquad i = 0, 1, ..., M. \tag{15}$$

These five equations will form the basis of the empirical steady state analysis later on in the paper.

4 Empirical Results

4.1 Data

We have data on labor inputs in 21 manufacturing industries from two sources: From the CPS we extracted data on industry number, usual weekly hours, usual weekly earnings, last grade attended, and completion of last grade attended (yes or no), for people who are employed as of the survey date. Then, to construct information on skilled labor with the cutoff of 12th grade, we classified all workers by the industry in which they work. A worker is defined as 'skilled' if his last grade attended is 13-or-higher or if he completed 12th grade.[7] Because of the relatively short sample period for the CPS (1979–1991), we also constructed a second data set for the years from 1960 to 1985 from 3 different sources: U.S. *National Income and Product Account, Handbook of Labor Statistics,* and *Survey of Current Business.*[8] For these data we do not observe educational attainments of workers, so we are forced to take the number of "production and non-supervisory" workers as a proxy for the number of "unskilled workers." While this is hardly a satisfactory division between skilled and unskilled, we find that this measure is strongly

[7]Other dividing points, such as junior college degree or college degree, yielded qualitatively similar but less precise estimates of the behavior of human capital.

[8]See appendix A for more details about the sources and construction of the data.

correlated with the measure based on the CPS data (the correlation coefficient for average skilled labor share by industry is 0.705), and the results are qualitatively similar for the periods in which the samples overlap.

From the raw data provided by these sources, we constructed data on total hours of skilled labor, total hours of unskilled labor, and the shares of inputs. For the shares of inputs we use both the cross-time averages of observed shares as well as well as the observed share for each period in each industry. The rationale for the using cross-time average is that we can smooth out a possible spurious correlation between observed shares and output. The exact formulae for the constructed data are also given in appendix A. Table 1 provides the key for industry indices that follow in the subsequent tables. Table 2 reports the summary statistics of the variables used in this study. Column 2-6 report the average growth rate of output, relative price, physical capital, total hours of skilled labor, and total hours of unskilled labor respectively. Column 7 reports average of Solow residuals modified to account for the change in skilled vs unskilled labor (hence MSR) for each industry, and finally the share of skilled labor is reported in column 8 of table 2.[9]

4.2 Estimation

In this section we ignore the steady–state restrictions and use both the time–series and cross–section variation. Given the general production function (1), standard calculations (assuming competition and CRS) yield

$$\Delta \ln y_{it} = \Delta \ln A_{it} + \alpha_{Kit} \Delta \ln K_{it} + \alpha_{sit} \Delta \ln(N^s_{it} H_t) + \alpha_{uit} \Delta \ln N^u_{it}, \qquad (16)$$

where the αs are share parameters; for example,

$$\alpha_{Sit} = F_2 N^s_{it} H_t / y_{it} \qquad (17)$$

or, simply, skilled labor's share in income.

We rewrite the log–differenced production relationship as

$$\Delta \ln y_{it} - (\alpha_{Kit} \Delta \ln K_{it} + \alpha_{sit} \Delta \ln N^s_{it} + \alpha_{uit} \Delta \ln N^u_{it}) = \alpha_{sit} \Delta \ln H_t + \Delta \ln A_{it}. \qquad (18)$$

[9]For formal definition of MSR see below.

The left–hand–side, which we will refer to as the "Modified Solow Residual" (MSR),[10] is observable, as is skilled labor's share on the right–hand–side. If, within a given time period, skilled labor's share is uncorrelated with $\Delta \ln A_{it}$, then a period–by–period regression will yield unbiased estimates of $\Delta \ln H_t$. In that case the mean of $\Delta \ln A_{it}$ across industries within a period is the "exogenous" contribution to the average MSR growth rate, while the remainder would be due to growth in skills. This would be equivalent to a pooled time series–cross section regression in which the $\{\Delta \ln A_{it}\}$ values are identified with the coefficients on the year dummies. If the mean of those coefficients were zero, it would indicate that all productivity growth was the result of growth in skills.

Of course as with any regression, a left–out factor that is correlated with α_{Sit} will get misattributed. But many such "left–out" factors are really precisely what we want to attribute to skilled labor. Anything that has an impact on productivity only to the extent the industry employs skilled labor is exactly what we are trying to capture in H.

4.2.1 Results

For the results using the CPS data, Figure 2a shows the estimated time path of $\log(H)$ and $\log(A)$. The estimated growth rates are integrated over time with the arbitrary initial values. One should bear in mind that the actual growth in productivity due to growth in H is not the growth in H itself, but that multiplied by skilled labor's share. Table 3 reports the main result of the paper, which is that growth in H makes a significant contribution to growth in productivity, while growth in A does not. This particular measure of H grows approximately 200% over the 10 year period, while generalized productivity A shows a 50% decrease. We tested two different hypotheses $H_0(H)$: growth rates of H are 0 for all t and $H_0(A)$: growth rates of A are 0 for all t. $H_0(H)$ is rejected at a 95% confidence level ($F\text{-}value = 4.062 > F(1, 198, \alpha = 0.05)$) while $H_0(A)$ is not rejected at a 95% confidence level ($F\text{-}value = 2.09 < F(1, 198, \alpha = 0.05)$). Thus the measured decline in A, while quantitatively large, is actually not statistically significant, while the growth in H is both large and significant.

[10]Note that the MSR differs from the conventional Solow residual to the extent that W^s differs from W^u, and $\Delta \ln N^s$ differs $\Delta \ln N^u$.

Since it is unlikely that human capital fluctuates so much on a yearly basis, we estimated (18) with a restriction that the growth rate of human capital are the same across time. In other word, $g_{H_t} = g_H \ \forall t$. The result is reported in columns 3 and 4 of Table 3. The result is essentially the same as the result of unrestricted regression, and we cannot reject the hypothesis of constant g_H over this sample.

The results are qualitatively similar for the longer period encompassed by the NIPA et al. data. These are presented in Figure 2b and Table 4. The estimated growth rates are integrated over time with the initial values based on results from regressions (not reported) in levels rather than first differences. The qualitative similarity of the results from the two data sets would suggest that the phenomenon of skilled labor–enhancing technological progress is not a phenomenon of the 1980s, but dates back at least to the 1960s. It should also be noted that the much–proclaimed productivity slowdown is not evident form the growth of H depicted in Figure 2b (which we must rely on because the CPS data do not cover the relevant sample period). This would suggest that the slowdown (to the extent it occurred in manufacturing) was not due to a slowdown in the growth of knowledge, but rather was due to factors that affect A, which could include increased regulation, mismeasurement of output, mismeasurement of capital, and so forth.

The estimated human capital shows steady growth until early 80s and then acceleration after that. Over the course of 25 years, human capital shows an average of 7.8% growth per year, which is much smaller than the CPS–based figure, though the CPS–based skilled labor share measures are larger (albeit covering a sample period that only partly overlaps). As indicated in Table 4, the null hypothesis that human capital has zero growth over the sample period is strongly rejected. (rejected at 1% significance level) As a result, human capital growth can explain approximately 65% of output growth of the U.S. manufacturing industry. Generalized productivity, on the other hand, shows only a little growth until 1973 or so, and slows down from that point on. On average, generalized productivity exhibits -0.74% annual growth, which it turns out is not significantly different from zero.

Even if one does not accept the production versus non–production distinction as a proxy for non–skilled versus skilled labor, the fact remains that it does account for productivity growth in our panel of 21 industries. This finding of what might be called "non–production worker–augmenting technical progress." is either a coincidence, or is the result of an important difference between the two types of workers. We would argue that the strong positive correlation between non–production workers' share and our CPS measure of skilled labor's share is the most natural explanation for our finding.

4.2.2 Steady State Estimation

We next examine the steady state implications of the model as embodied in (13)–(15). Some of the implications are not specific to the model but are of interest more generally in characterizing productivity growth and sectoral trends in the data. It turns out that if the production function is not Cobb–Douglas, the steady–state relationships can provide a second, independent measure of the relative importance of $\hat{A}$ and $\hat{H}$. This can be done via estimation or calibration. The results provide further confirmation of the non–steady–state findings that $\hat{H}$ is the only significant contributor to productivity growth.

Calibration can determine the parameter θ and steady–state values of $\hat{A}$ and $\hat{H}$ as follows: First, note that if we substitute the unskilled labor equation into the capital equation from (13), we get

$$\hat{W}^u + \theta(\hat{N}^u - \hat{K}) = 0. \tag{19}$$

Using the aggregate manufacturing values of $\hat{W}^u$ and $\hat{N}^u - \hat{K}$, we can obtain a calibrated value for θ. We can similarly obtain values for $\hat{A}$ and $\hat{H}$ using the skilled labor equation and one of the other equations in (13), again provided our estimate of theta differs sufficiently from one. The result of this exercise is provided in the first panel of Table 5 for the two data sets. Although there are no standard errors for this exercise, the results are qualitatively similar to the non–steady–state results insofar as the growth of H is considerably larger than the growth of A, though the difference in contribution to growth (after factoring in skilled labor's share, which is 0.298 in the NIPA data set and 0.620 in the CPS data set for aggregate manufacturing) is not as large. Perhaps the more important thing is that the calibrated value of θ is considerably less than one, which allows us to separate out the effects of H and A at all.

Given the likelihood that $\theta < 1$, we can go on to estimate the system (13), using non–linear least squares and imposing cross–equation restrictions to estimate the three unknown values. For this purpose we use the aggregate manufacturing wage growth numbers for skilled and unskilled labor for $\hat{W}^s + \hat{H}$ and $\hat{W}^u$ respectively (since the model implies common wage rates for all industries), but similar results obtained using each industry's own computed wage growth series. These results are presented in the second panel of Table 5. They are consistent with the calibration results, and show that the growth in H is statistically significant, at least in the NIPA data, whereas the growth in A is not. Also, the estimates of θ differ significantly from one. The table also gives the R^2 values for the three equations as indicated.

Finally, we can estimate (14) and (15) to get estimates of γ and another set of estimate of $\hat{H}$ and $\hat{A}$. These results are shown in the bottom panel of Table 5. The estimates of γ, the elasticity of substitution in utility across goods, do not differ significantly from one (though very nearly does for the CPS data). The estimates of $\hat{H}$ and $\hat{A}$ are not surprisingly qualitatively similar to those from the non–steady–state analysis, though it is somewhat disturbing that the estimate of $\hat{A}$ from the CPS data is significantly negative in this case.

One question that naturally arises is the interpretation of the error terms in these equations. Certainly $\hat{A}_i$, the idiosyncratic or industry–specific general productivity growth factor, is one component of the error. But this is common to all three equations. Any other disturbances would have to be due to measurement error of some sort. But similar results were also obtained from other estimation techniques, specifically equation–by–equation OLS and Two–Stage Least Squares, which suggests that the results are reasonably robust.

Overall the steady state results provide additional independent evidence that skilled–labor–augmenting technical change is the primary contributor to productivity growth, and also provide estimates of other parameters of the model that are plausible, suggesting that it is a useful framework for analysis of these data.

4.3 Discussion: Externalities

We would argue that the results described above cast some doubt on the importance of production externalities. To make the argument more precise, we extend the model to allow for such externalities. Now suppose a representative firm in industry i has the following technology:

$$Y_{it} = A_{it}F_i(K_{it}, N^s_{it}H_t, N^u_{it}, X_t), \tag{20}$$

where X_t is a vector of aggregate variables (some of which may be measurable) exogenous to the firm. X_t could include (along the lines of Lucas, Romer), the average skill level in the economy as a whole: or the average skill level of skilled workers in the economy as a whole. Assuming competition and CRS, we have

$$\Delta \ln Y_{it} = \Delta \ln A_{it} + \alpha_{Kit}\Delta \ln K_{it} + \alpha_{Sit}\Delta \ln(N^s_{it}H_t) + \alpha_{Uit}\Delta \ln N^u_{it} + \alpha'_{X\,it}\Delta \ln X_t, \tag{21}$$

where, again,

$$\alpha_{Sit} = F_2 N^s_{it}H_t/Y_{it} \tag{22}$$

or, simply, skilled labor's share in income, and other share parameters are defined analogously.

Leaving aside the identity of X for the moment, we rewrite the log–differenced production relationship as

$$\Delta \ln Y_{it} - (\alpha_{Kit}\Delta \ln K_{it} + \alpha_{Sit}\Delta \ln N^s_{it} + \alpha_{Uit}\Delta \ln N^u_{it}) = \alpha_{Sit}\Delta \ln H_t + \epsilon_{it}. \tag{23}$$

where $\epsilon_{it} \equiv \alpha'_{X\,it}\Delta \ln X_t + \ln A_{it}$. The left–hand–side, which we will refer to as the "Modified Solow Residual (MSR)", is observable, as is skilled labor's share on the right–hand–side. If, within a given time period, skilled labor's share is uncorrelated with ϵ_{it}, then a period–by–period regression will yield unbiased estimates of $\Delta \ln H_t$. In that case the mean of ϵ_{it} across industries within a period is the "exogenous" contribution to the average MSR growth rate, while the remainder would be due to growth in skills. If under those assumptions that mean were zero, it would indicate that all productivity growth was the result of growth in skills.

The big question, of course, is what is in ϵ_{it}, and how likely is it to be uncorrelated with α_{Sit}? Throughout we will assume with essentially no loss in generality that A_{it} is an exogenous stochastic process uncorrelated with α_{Sit}. (Correlated factors can be considered part of X.) We consider several cases in which this assumption is valid:

1. X_t is a constant

2. α_{Xit} does not depend on i

3. α_{Xit} is orthogonal to α_{Sit}, and either A_{it} is constant or X_t is uncorrelated with H_t

The first case would correspond to the simplest generalization of the neoclassical growth model to include human capital accumulation. Again, a test of the hypothesis that the mean of ϵ_{it} is zero is a test of zero exogenous growth. The second case would be consistent with the presence of any externality that entered the production technologies of all industries symmetrically, for example, $Y_{it} = A_{it}F_i(K_{it}, N^s_{it}H_t, N^u_{it})X_t$. This would include the case of Lucas-Romer externalities in aggregate skill levels. The third case allows X to enter differently in different industries, but in a way that leaves the overall effect uncorrelated with skilled labor's share.

Under any of these assumptions, we identify the relative contribution of growth in H to productivity growth. Our finding is that contribution is essentially 100 percent. Are the assumptions plausible? Yes. Are there models that would violate these assumptions that are also plausible? Yes. For example, suppose we have

$$Y_{it} = A_{it}F_i(K_{it}, N^s_{it}H_t, N^u_{it})X_t^{\delta_i} \tag{24}$$

where δ_i is correlated with industry i's skilled labor share, and where X_t is correlated with, or even equal to, H_t. Then there is no way this exercise can sort out the internalized contribution of H from the external contribution. In principle, however, if X is observable and quantifiable, it can be entered into the regression equation to yield estimates of α_X. This is a subject for future research.

Note that the results have no bearing on the presence or absence of externalities in human capital accumulation. We have not modelled the process by which H is determined for individuals, industries, or in the aggregate.

5 Concluding Remark

Identifying the sources of economic growth has long been a goal of empirical research into economic growth. Although many have suspected that human capital may be the single most important factor, the evidence has been inconclusive largely due to the inherent difficulty of measuring human capital and identifying its contribution separately from other factors. The contribution of this paper is to link human capital growth to the presence of skilled labor in an industry, and thereby to identify separately its role from general growth in productivity. Given this interpretation, the results confirm the importance of human capital as the prime factor in growth. Moreover, they leave essentially no role for growth in total factor productivity unlinked to skilled labor.

Perhaps the most interesting finding is the qualitative similarity of the results from the year–by–year estimation of human capital growth with the results from the steady–state analysis. We would argue that the two findings represent different pieces of evidence for the same phenomenon, because the steady–state findings are not a logical implication of the year–by–year results. Indeed human capital growth is not even identified in the steady–state analysis unless the production function has an elasticity of substitution that differs from one.

These results are closely related to the recent literature on the growth in the wage gap between skilled and unskilled labor. It should be emphasized, however, that growth in "knowledge" as identified in the empirical work does not imply anything about growth in the earnings gap between skilled and unskilled labor. That gap depends, of course, on supply as well as demand. Absent changes on the supply side, skilled labor augmenting technical change can easily lead to offsetting changes in the wage rates for the two worker types per unit of effective labor, so that the earnings gap remains the same.

A Data

Information regarding the source of raw data and the exact formulation of constructed variables are provided.

Source:
All the data used in this study come from the following three sources.

1. Data set from Shapiro(1987): NIPA and Survey of Current Business[11]

2. BLS data: Handbook of Labor Statistics (current employment statistics program) by Bureau of Labor Statistics.

[11] Authors thank Bob King for making this data available.

3. CPS data: Abstracts of Current Population Survey (Outgoing Rotation Group) by Bureau of Census.

Raw data:

1. Y: Real Output in 1982$ (NIPA table 6.2)
2. Y^N: Nominal Output (NIPA table 6.1)
3. N: Total hours employed (NIPA table 6.11)
4. Comp: Total compensation for labor (NIPA table 6.4)
5. W&S: Total wage and salary (NIPA table 6.5)
6. FTE: Full-time-equivalent employees (NIPA table 6.7)
7. K: Net physical capital in 1982 dollars (see, for example, Survey of Current Business august 1986)
8. Emp: Total number of Employees (BLS Handbook table 70-71)
9. P&NS: Total number of production workers (BLS Handbook table 70-71)
10. W^u: Average hourly earnings of production workers (BLS Handbook table 81)

Constructed Variables Using BLS data:

The following are the precise definitions of transformed variables constructed using the BLS data. Here, classification of labor skill is made by the kind of work workers do.

- $P_i = \left(Y_i^N \div Y_i\right) \div \left(Y_{mfg}^N \div Y_{mfg}\right)$ where subscript $_{mfg}$ stands for aggregate manufacturing industry.
- $N_i^u = N_i \times (\text{P\&NS}_i \div \text{Emp}_i)$
- $N_i^s = N_i - N_i^u$
- $S_{K_i} = 1 - \frac{1}{T}\sum_{t=1}^{T}(Comp \div Y^N)_{it}$

- $\delta_i = \frac{1}{T}\sum_{t=1}^{T}(W^u \cdot N^u \div \text{W\&S})_{it}$
- $S_{N_i^u} = \delta_i \times (1 - S_{K_i})$
- $S_{N_i^s} = (1 - \delta_i) \times (1 - S_{K_i})$

BLS data are yearly observations, and cover 21 U.S. manufacturing industries from 1960 to 1985.

Constructed Variables Using CPS data:
Below are the definitions of hours of skilled work employment and skilled labor share when the CPS data are used. Here, classification of labor skill is made by years of education, not by the kind of work they do. Skilled labor employed for each industry, or N_{it}^s, is defined as,

$$N_{it}^s = \frac{(\text{total usual weekly hours of skilled workers})_{it}}{(\text{total usual weekly hours of both skilled and unskilled workers})_{it}} \times N_{it}$$

where the last term N_{it} is defined as 'total hours employed', and is obtained from NIPA. Of course, $N_{it}^u = N_{it} - N_{it}^s$. Similarly, δ, or skilled labor share, is defined as

$$\delta_{it} = \frac{(\text{total usual weekly earnings of skilled workers})_{it}}{(\text{total usual weekly earnings of both skilled and unskilled workers})_{it}}$$

The CPS data are also yearly, and cover 21 U.S. manufacturing industries from 1979 to 1991.

References

[1] Barro, Robert, "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics* 106 (May 1991) 407-444.

[2] Barro, Robert and Xavier Xala-i-Martin, "Economic Growth and Convergence Across United States", mimeo, Harvard University, (September 1989).

[3] Benhabib, Jess and Boyan Jovanovic, "Externality and Growth Accounting", *American Economic Review* 81 (March 1991), 82-113.

[4] Caballero, Ricardo and Richard Lyons, "The Role of External Economies in U.S. Manufacturing", *NBER* WP No. 3033, (July 1989).

[5] Feenstra, Robert C., James R. Markusen, and William Zeille, "Accounting for Growth with New Inputs: Theory and Evidence", *American Economic Review* 82 (May 1992), 415-421.

[6] Gort, Michael and Richard A. Wall, "Obsolescence, Input Augmentation, and Growth Accounting", mimeo, State University of New York at Buffalo, (July 1993).

[7] Hulten, Charles, "Growth Accounting When Technical Change Is Embodied in Capital", *American Economic Review* 82 (September 1992), 964-980.

[8] Levine, Ross and David Renelt, "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review* 82 (September 1992), 942-963.

[9] Lucas, Robert E., "On the Mechanics of Economics Growth", *Journal of Monetary Economics,* 22 (1988), 3-42.

[10] Romer, Paul M., "Crazy Explanation for the Productivity Slowdown," in Stanley Fischer, ed., *NBER Macroeconomic Annual*, Cambridge, MA: MIT Press, 1987, 163-202.

[11] Romer, Paul M., "Two Strategies for Economic Development: Using Ideas and Producing Ideas," *mimeo* (May 1992).

[12] Scott, Maurice FG., "Explaining Economic Growth", *American Economic Review* 83 (May 1993), 421-425.

[13] Wolff, Edward N., "Capital Formation and Productivity Convergence Over the Long Term", *American Economic Review* 81 (June 1991), 565-579.

Figure 1a

Conventional Solow Residuals on Skilled Labor Share

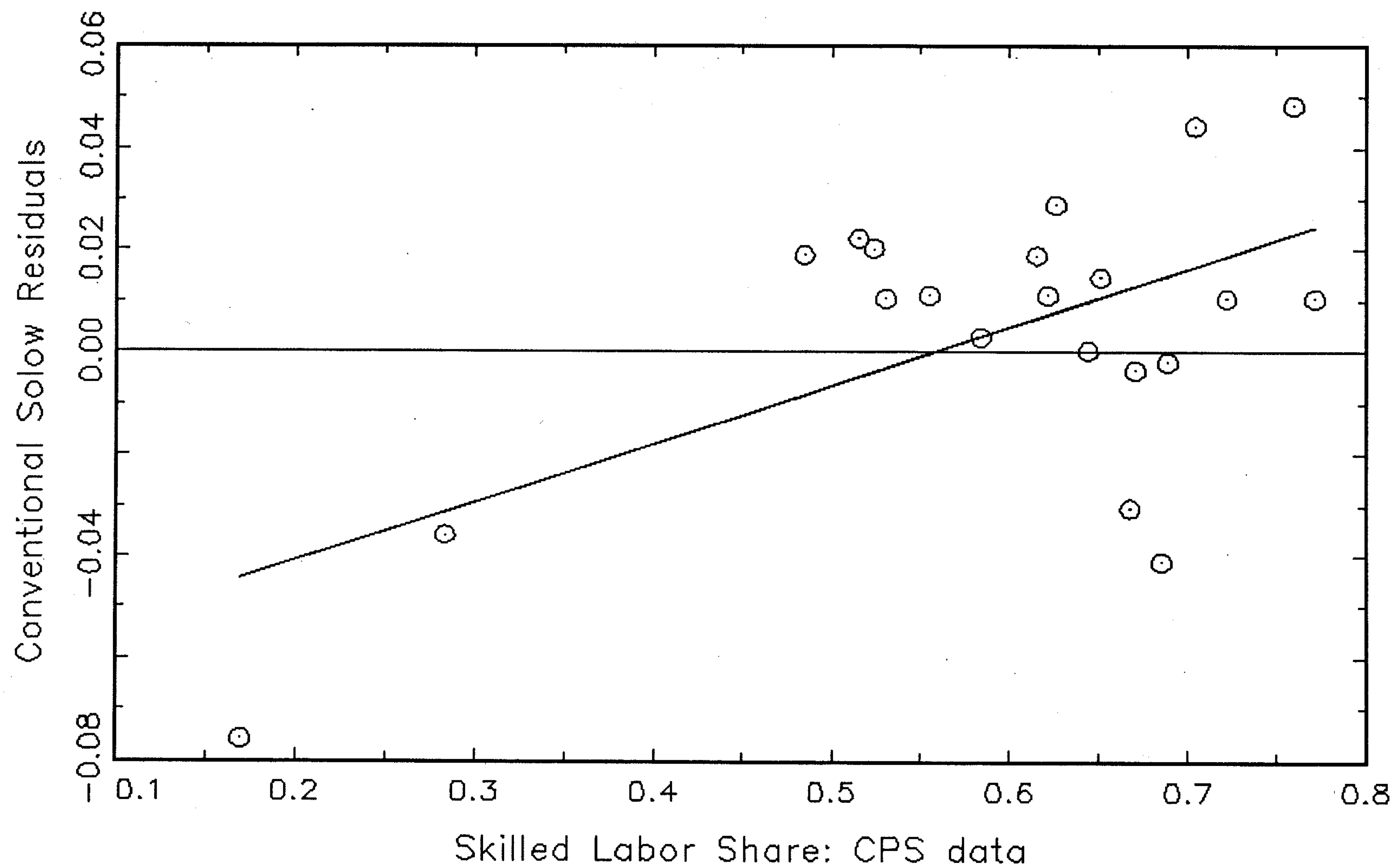

Figure 1b
Conventional Solow Residuals on Skilled Labor Share

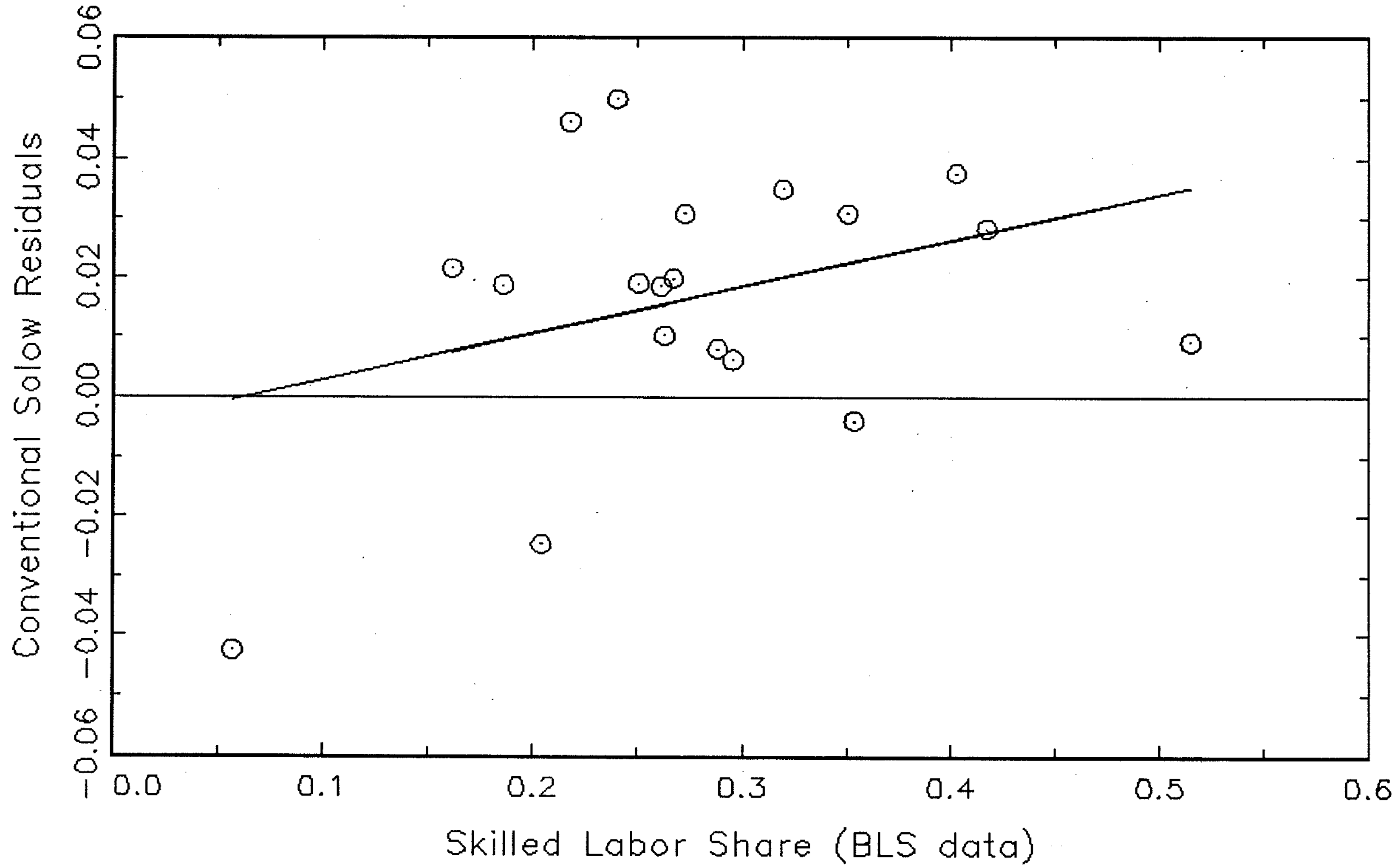

Figure 2a

Estimated Human Capital and Generalized Productivity

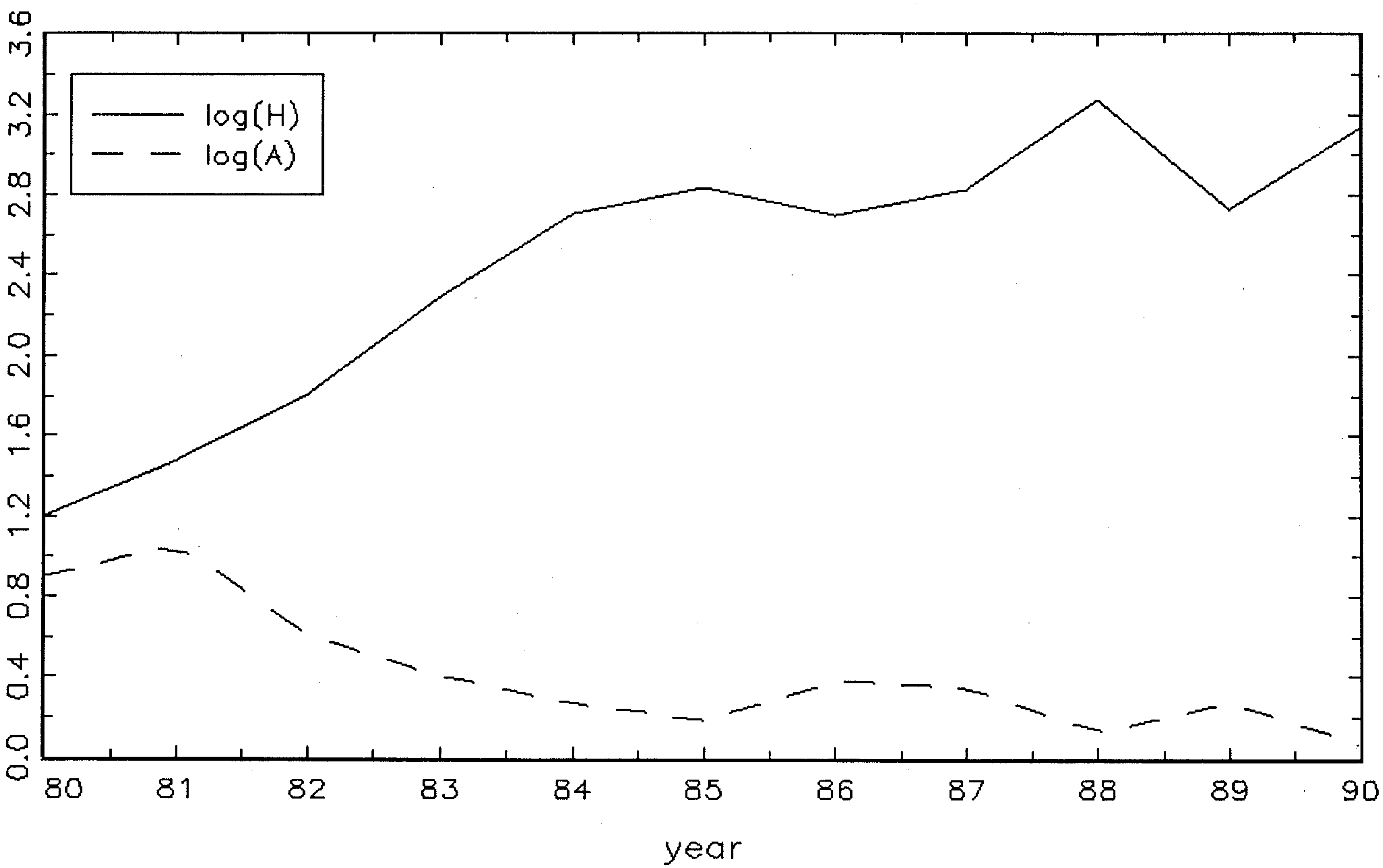

Figure 2b
Estimated Human Capital and Generalized Productivity

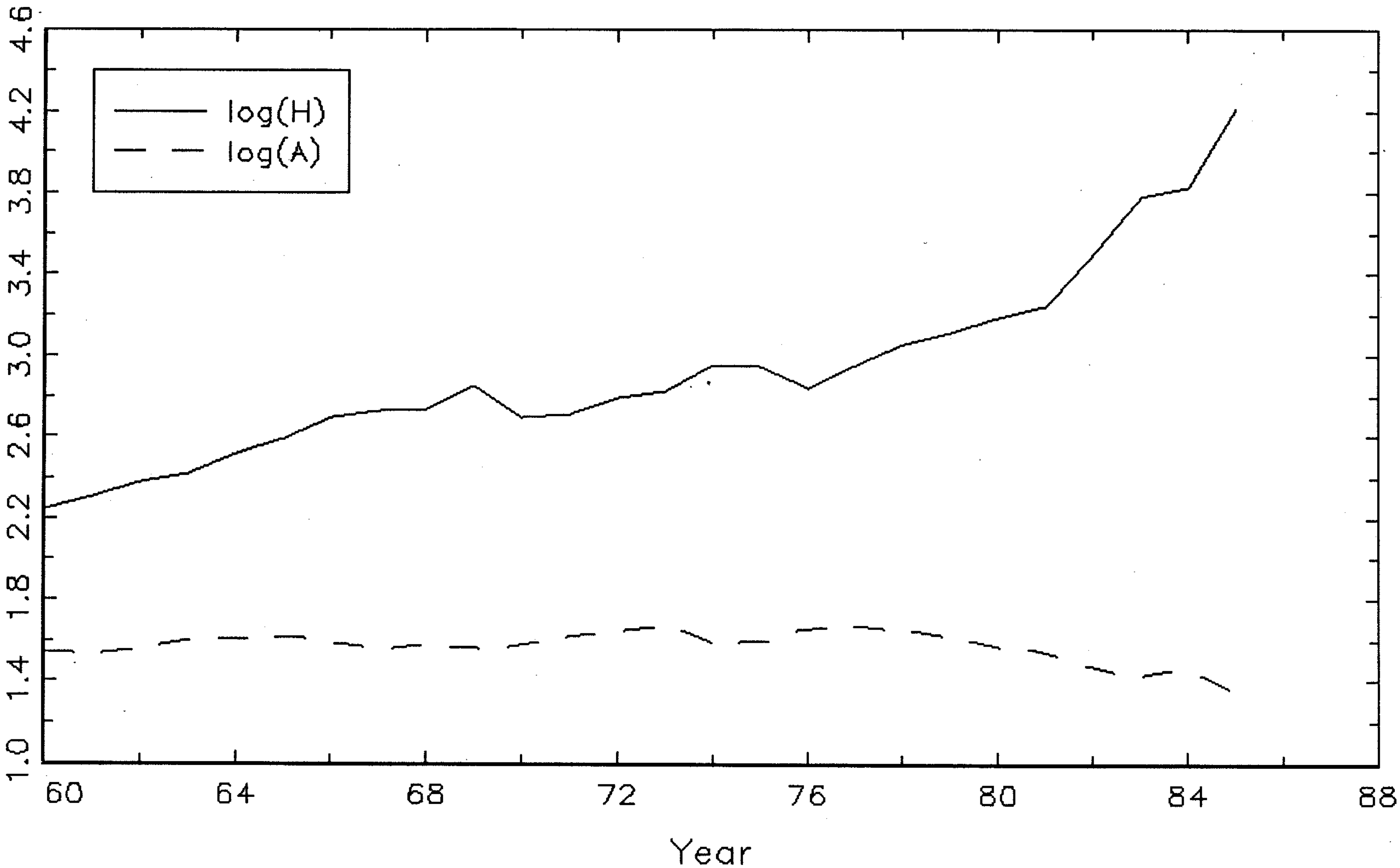

Table 1

Key to industry index

Industry index*	Industry Name
15	Lumber and wood products
16	Furniture and fixtures
17	Stone, clay, and glass products
18	Primary metal industries
19	Fabricated metal products
20	Machinery except electrical
21	Electric and electronic equipment
22	Motor vehicles and equipment
23	Other transportation equipment
24	Instrument and related products
25	Miscellaneous manufacturing industries
27	Food and kindred products
28	Tobacco manufactures
29	Textile mill products
30	Apparel and other textile products
31	Paper and allied products
32	Printing and publishing
33	Chemicals and allied products
34	Petroleum and coal products
35	Rubber and miscellaneous plastic products
36	Leather and leather products

* These numbers correspond to the line numbers of *Survey of Current Business* and *National Income and Product Account.*

Table 2

Average Growth rate of Variables (from 1960 to 1985) (NIPA data)

Industry*	Y	P	K	Ls**	Lu***	MSR+	S(Ls)++
15	3.45%	0.11%	2.30%	1.67%	-0.11%	2.42%	0.1614
16	3.13%	0.71%	3.49%	1.76%	0.91%	1.43%	0.2629
17	1.99%	0.61%	1.74%	0.74%	-0.38%	1.47%	0.2505
18	-0.40%	1.13%	1.95%	-0.52%	-1.67%	0.04%	0.2469
19	2.48%	1.06%	3.57%	1.02%	0.45%	1.12%	0.2955
20	5.49%	-1.11%	4.67%	2.62%	0.93%	2.91%	0.3505
21	6.04%	-1.78%	6.25%	2.52%	1.12%	3.20%	0.4027
22	3.67%	-1.13%	3.14%	0.83%	0.77%	1.95%	0.2400
23	2.47%	1.46%	3.73%	1.28%	-0.04%	1.54%	0.5149
24	4.87%	-0.51%	5.55%	3.16%	1.46%	1.57%	0.4175
25	2.05%	0.65%	3.08%	1.15%	-0.68%	1.15%	0.2881
27	2.65%	-0.34%	2.13%	-0.68%	-0.31%	2.04%	0.2671
28	-0.34%	2.38%	5.53%	1.48%	-2.16%	-4.85%	0.0571
29	3.80%	-2.36%	1.47%	0.18%	-1.32%	4.15%	0.2175
30	2.33%	-0.28%	4.17%	0.99%	-0.63%	1.88%	0.1859
31	3.60%	0.25%	3.45%	1.22%	0.27%	1.97%	0.2611
32	2.68%	1.78%	4.05%	2.60%	1.09%	0.20%	0.3534
33	5.08%	-1.24%	3.90%	1.52%	0.53%	2.70%	0.3194
34	1.64%	2.87%	2.88%	-0.17%	-0.92%	0.08%	0.2039
35	5.38%	-0.95%	4.56%	2.60%	2.61%	2.14%	0.2725
36	-0.99%	0.13%	1.80%	-1.52%	-3.24%	1.12%	0.2204

* Numbers correspond to the line number of the *Survey of Current Business*. Key to these number can be found in table 1.

** Skilled labor hour

*** Unskilled labor hour

\+ Modified Solow residuals

++ Skilled labor share

For the exact definition of these variables, see appendix A.

Table 2 (cont.)

Average Growth rate of Variables (Using CPS Data: 1979-1990)

Industry*	Y	P	K	Ls**	Lu***	MSR+	S(Ls)++
15	1.58%	-0.87%	-0.63%	6.01%	-1.95%	-1.04%	0.5152
16	0.87%	1.90%	3.73%	-0.39%	0.71%	0.20%	0.6440
17	-0.29%	-0.55%	-1.46%	-14.05%	-0.20%	9.54%	0.6705
18	-4.52%	1.06%	-1.76%	-3.59%	-4.54%	-1.28%	0.6851
19	0.49%	0.20%	2.52%	-0.57%	-1.53%	0.32%	0.6887
20	6.39%	-5.33%	9.43%	0.05%	-1.65%	4.44%	0.7042
21	2.71%	-1.26%	6.06%	-10.05%	-0.79%	8.70%	0.7220
22	1.75%	0.64%	3.33%	-3.89%	0.43%	4.22%	0.7716
24	7.40%	0.68%	11.01%	-9.50%	6.44%	12.38%	0.7596
25	3.02%	0.68%	1.93%	-4.15%	0.00%	4.52%	0.5239
27	2.28%	0.80%	2.49%	-8.27%	0.63%	5.84%	0.5549
28	-8.56%	13.58%	-1.26%	-2.51%	-2.56%	-7.05%	0.1693
29	1.18%	-0.67%	-0.09%	6.07%	-4.50%	-1.79%	0.6154
30	1.45%	-0.63%	1.97%	2.19%	-3.18%	0.08%	0.6508
31	2.47%	1.16%	3.24%	-2.26%	0.86%	2.42%	0.5307
32	1.69%	3.57%	3.85%	-1.40%	3.60%	0.94%	0.5840
33	3.10%	1.52%	2.51%	-1.38%	0.16%	2.69%	0.4848
34	0.08%	4.91%	5.20%	-9.15%	-0.65%	-0.73%	0.2837
35	4.63%	-1.56%	4.26%	-4.57%	3.86%	5.97%	0.6258
36	-1.81%	0.85%	0.42%	-16.85%	-1.43%	9.56%	0.6672
Mfg.	2.31%	0.00%	3.48%	-3.58%	-0.04%	3.52%	0.6213

* Numbers correspond to the line number of the *Survey of Current Business*. Key to these number can be found in table 1.

** Skilled labor hour

*** Unskilled labor hour

\+ Modified Solow residuals

++ Skilled labor share

For the exact definition of these variables, see appendix A.

Table 3

Regression Equation: $\Delta MSR_{it} = \beta_{0t} + \beta_{1t} \cdot (\text{skilled labor share})_i + \varepsilon_{it}$

Restriction	None		$g_H(t)=g_H$ for all t	
Variable	H	GP	H	GP
Avg. yearly growth Rate	19.4%	-8.5%	19.4%	-8.5%
F-statistic*	4.062**	2.090	4.080**	2.099
R^2	0.366		0.337	
Adj. R^2	0.292		0.267	

* H_0: sum of all $g_{Ht} = 0$ and sum of all $g_{GPt} = 0$ respectively.

** Reject the null (of insignificant regression) at the 99% confidence level.

CPS household survey data are used for this regression.

Table 4

Regression Equation: $\Delta MSR_{it} = \beta_{0t} + \beta_{1t} \cdot (\text{skilled labor share})_i + \varepsilon_{it}$

Restriction	None		$g_H(t)=g_H$ for all t	
Variable	H	GP	H	GP
Avg. yearly growth Rate	7.8%	-0.74%	7.9%	-0.74%
F-statistic*	9.506**		9.641**	
R^2	0.276		0.251	
Adj. R^2	0.202		0.212	

* H_0: sum of all $g_{Ht} = 0$ and sum of all $g_{GPt} = 0$ respectively.

** Reject the null (of insignificant regression) at the 99% confidence level.

BLS production/non-production workers data are used for this regression.

Table 5

Calibration of Steady State

	θ	$\hat{A}(\%)$	$\hat{H}(\%)$
NIPA	0.675	0.349	3.159
CPS	0.420	0.843	2.000

Steady State Estimation Results I

	θ	$\hat{A}$	$\hat{H}$	R^2 (K, N^u, N^s)
NIPA	0.644	0.673	2.417	0.435, 0.542, 0.590
	(0.057)	(0.412)	(0.762)	
CPS	0.737	-1.679	3.257	0.318, 0.380, 0.630
	(0.104)	(1.871)	(2.891)	

Steady State Estimation Results II

	γ^{-1}	$\hat{A}$	$\hat{H}$	R^2 (y, MSR)
NIPA	0.958	-0.746	7.896	0.453, 0.194
	(0.241)	(1.082)	(3.696)	
CPS	0.698	-5.753	10.901	0.517, 0.393
	(0.159)	(1.958)	(3.195)	

Note: Standard errors are in parentheses.